1965

Using dynamic programming and Markov processes to solve a multiple delivery inventory problem

Fred W. Ehrman
Lehigh University

USING DYNAMIC PROGRAMMING AND MARKOV PROCESSES TO SOLVE A MULTIPLE DELIVERY INVENTORY PROBLEM

by

Fred W. Ehrman

A THESIS

Presented to the Graduate Faculty

of Lehigh University

In Candidacy for the Degree of

Master of Science

Lehigh University

1965

This thesis is accepted and approved in partial fulfillment of the requirements for the degree of Master of Science.

May 20, 1965
Date

John M. Carroll
Professor in Charge

Head of the Department

ACKNOWLEDGEMENTS

The author expresses his appreciation to Professor John M. Carroll for his overall guidance and helpful criticism; and to Professor William A. Smith, Jr. and Dr. N. P. Loomba for their helpful criticism.

Above all, I thank Mr. Ken E. Larson for his counsel on the initial approach to the problem and his counsel on the evaluation of the results.

TABLE OF CONTENTS

TABLE OF CONTENTS (cont'd)

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

A more economic inventory policy can be obtained by specifying the delivery schedule properly. To determine such a policy requires a great amount of computing time. However, this computing time can be reduced with a dynamic programming and markov processes approach. The thesis will cover the procedure in determining an optimal policy using such an approach.

1. INTRODUCTION

Most inventory models that have a solution place an order for a single delivery at a later date. For example, at the beginning of the month an order is determined and placed for delivery at the end of the month. However, deliveries in some cases may be made more often at less cost. For example, if one determines and places an order at the beginning of the month, it may be more economical to specify partial delivery at the end of each week during the month. The problem now becomes one of determining a specific delivery to be made at the end of each of four weeks.

The solution to this type of problem can be found by enumerating all possible combinations of delivery schedules, computing the cost of each, and finally selecting the combination that yields the greatest economy. To make such a procedure possible one must establish a manageable number of states or levels of inventory that the system can take on. For example, the states can be defined as in Table I with the corresponding class marks and class intervals. In this example the delivery amounts are in increments of 100. If for this problem 4

Table I. Example Of System With 6 States

State No.	Class Mark	Class Interval
1	-200	$-\infty$ to -151
2	-100	-150 to -51
3	0	-50 to 50
4	100	51 to 150
5	200	151 to 250
6	300	251 to ∞

optional delivery amounts are considered for each of 4 weeks, 4 x 4 x 4 x 4 or 256 policies must be evaluated for each state. Now if all 6 states are included in one system, 256^6 or 281, 474, 976, 710, 565 policies must be evaluated. Since the effort is too great in evaluating so many policies, another method must be used. Dynamic programming [2, 3, 4, 19] has been developed so that not all policies need to be evaluated. Ronald Howard [11] goes further and takes into account the probability of transition from one state to another by combining markov processes with dynamic programming. In chapters 4 and 5 dynamic programming and markov processes are discussed in greater detail and in chapter 6 the application of this method to the solution of the problem is discussed.

2. DESCRIPTION OF THE PROBLEM

The inventory system that will be discussed is one where the product has a continuous demand and is continuously produced. The question that remains is what quantity should be ordered for delivery. So that the problem can be described more clearly, a specific procedure will be discussed. This procedure is one that is in operation presently at a manufacturing location.

Present Procedure

At the beginning of the month the requirements for the month are determined. The inventory level is checked and a forecast of the months demand made. If the inventory level is above or below the desired level, a correction is made such that the correction will be made one-third of the way by the end of the month. This inventory model is shown in Figure 1. The computation of the order quantity is as follows:

$$Q = FD + 1/3\ (DL - PL)$$

where:

Q = Order Quantity

FD = Forecasted Demand

DL = Desired Inventory Level

PL = Present Inventory Level

The computed requirement is then placed as an order to the supplier which may or may not be the company's own manufacturing shop. Partial delivery is made each week. The amount will depend on what has been completed and inspected. The quantities that can be expected to be delivered are given in percent of order quantity in Table II.

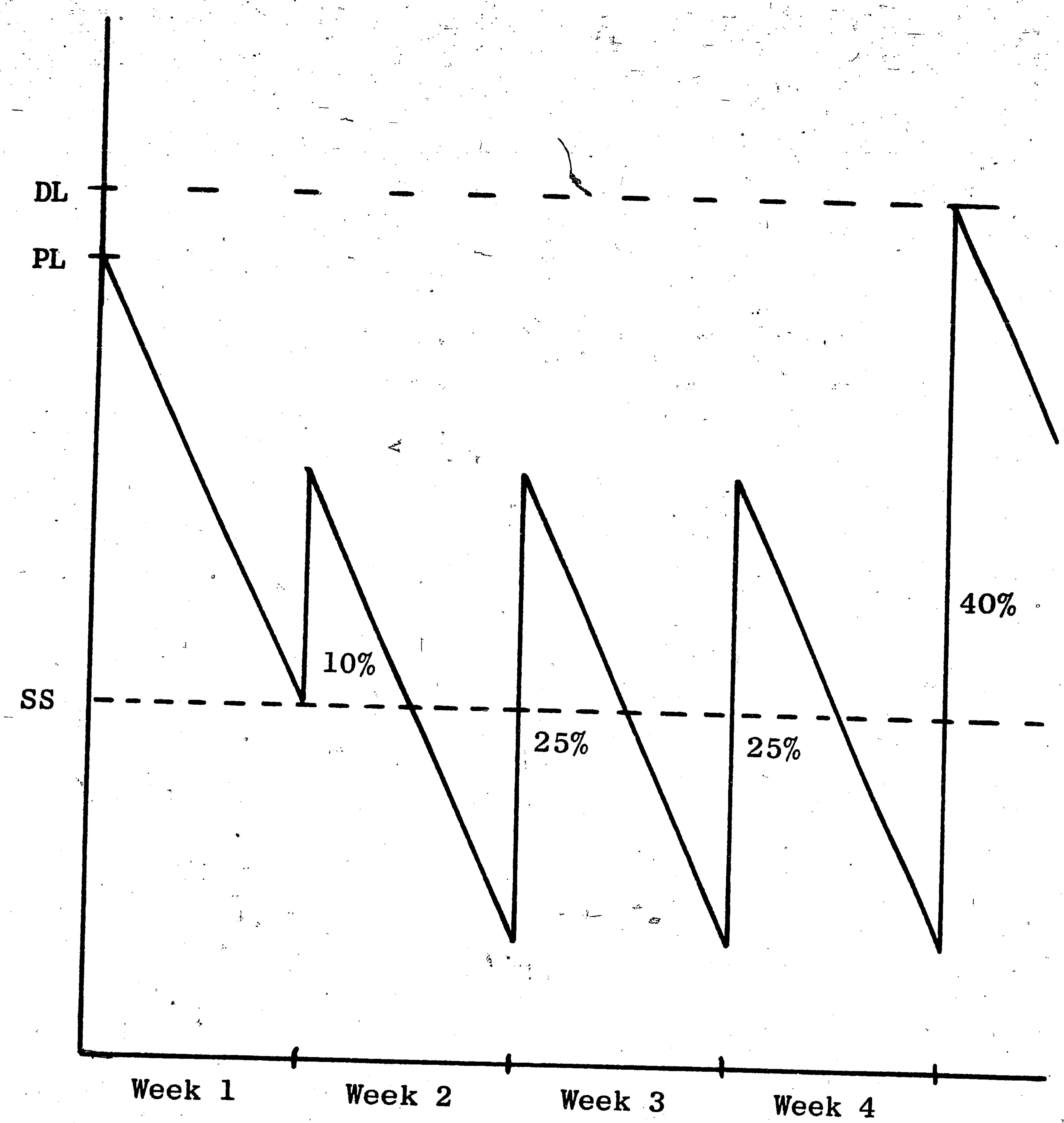

Figure 1

Present Inventory Model

Table II. Expected Delivery Quantities

Week	% Delivered
1	10
2	25
3	25
4	40

Advantages of Present Procedure

The present procedure is simple and once established will require comparatively little computation time. After the desired inventory level has been established, the computation for order quantity can be done on a desk calculator. Or, if many different products are ordered, a computer can be used.

If the order is to the Company's own shop, another advantage is that the shop has the freedom to schedule the production. For example, the shop can schedule to produce a product any time during the month which may mean producing the entire order during the last week.

Disadvantages of Present Procedure

Since the order quantity is the requirement for the month and the supplier is free to make delivery or partial deliveries anytime during the month, the desired inventory level must be set higher. The higher level will insure a lower probability of stockout.

Proposed Procedure

In this procedure an ordering policy is established. That is, after the beginning of the month inventory level is checked, an amount is specified for delivery during each of the four weeks. It is in this procedure that a manageable number of states or inventory levels as illustrated in Table I must be established.

Advantages of Proposed Procedure

By specifying delivery quantities inventory is controlled within closer limits. The major adjustment is during the first week after placing the order and therefore the inventory level is brought into line quickly. Because the response is quicker the probability of stockout is reduced while keeping the average inventory level lower which results in lower holding costs.

Another advantage is that once a policy is established no computations need to be made when determining an order and delivery schedule. One merely ascertains the inventory level and then looks up the ordering policy for that inventory level and places the order.

Disadvantages of Proposed Procedure

The computing time required to determine a policy is large. It is necessary to use a higher speed computer such as an IBM 7090 so that the time does not become excessive. Another disadvantage is that this procedure puts restrictions on the supplier. The supplier no longer has the freedom he had before when he could schedule the entire production during the last week if he so desired. As a result if the company's own manufacturing shop is the supplier, the shop may be forced to operate at a non-optimal level. Also the major adjustment of production is during the first week of the month. Less time is available to plan any manpower smoothing that the shop is doing.

3. INVENTORY MODEL

To discuss the inventory model, states will be defined as given in Table I. The model shown in Figure 2 is set up using these six states. Also, the demand distribution is defined as follows:

P_1 = Probability of no demand

P_2 = Probability that 100 are demanded

P_3 = Probability that 200 are demanded

P_4 = Probability that 300 are demanded

Assumptions

The following assumptions are made in the proposed inventory model:

1. The process is ergodic. That is, the weekly demand distribution function does not change with time. The discrete demand probabilities will be the same in the future as now.

2. The demand distribution is discrete rather than continuous.

3. The order quantity will be determined after the last delivery of the preceding month, and the order will be placed at the beginning of the current month.

4. Delivery~~, if~~ any, will be made at the end of each week.

Proposed Inventory Model

The diagram in Figure 2 shows the initial inventory level at 100 units (State 4). The inventory system at the end of week 1 will be in state T1(7) with probability P_1, state T1(6) with probability P_2, state T1(5) with probability P_3 and state T1(4) with probability P_4. At this point a delivery n_1 is made which in this diagram is 300 units. Therefore, the inventory level is raised 3 states. The diagram shows only partially the changes so that the process may be seen more clearly. In

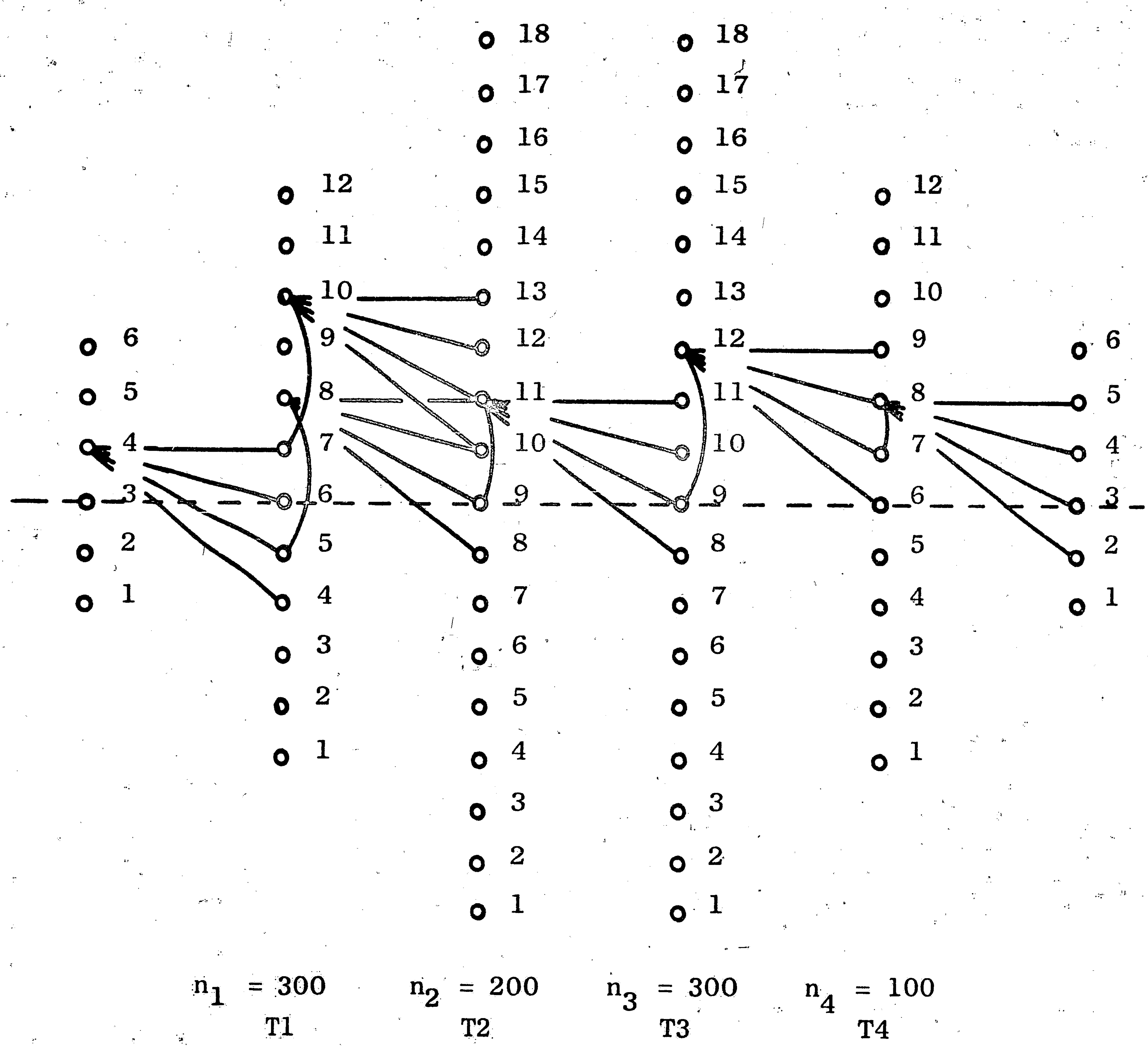

Figure 2

Diagram Showing Partially The Possible Inventory Levels (States) For The Next 5 Periods Given An Initial Level And Delivery Schedule.

actuality the inventory level changes from state T1(7) to state T1(10), from state T1(6) to state T1(9), and so on until all changes are accounted for. In each change the appropriate probability is carried along.

During the second week the number of possible transitions increases. The probability of being in state T2(13) at the end of week 2 before delivery N_2 is P_1P_1. That is, the probability of being in state T1(10) multiplied by the probability of no demand. Also, the probability of being in state T2(12) before delivery N_2 is $P_2P_1 + P_1P_2$. That is, the probability of being in state T1(9) multiplied by the probability of a 100 unit demand plus the probability of being in state T1(10) multiplied by the probability of no demand. This procedure is continued until all probabilities are computed.

At the end of week 2 a delivery n_2 is made and all of the inventory levels with their respective probabilities must be changed in the same manner as at the end of week 1. This process is continued until the end of week 5 is reached.

The reason that the process must be taken to the end of week 5 is that the events during week 5 are chargeable as far as economics is concerned to the decision being made at the beginning of week 1. Also, the events during week 1 are not chargeable to the present decision. The charges to the present decision are not started until the first delivery is made. This delivery is the first response to the decision.

4. DYNAMIC PROGRAMMING [2, 3]

Since it takes a great amount of computing time to determine the cost of all possible combinations of deliveries, it becomes necessary to find another way to find an optimal solution. This can be done by using dynamic programming. Dynamic programming is based on the use of functional equations and the principle of optimality keeping in mind computation techniques that can be used on digital computers. Usually, this technique is an iterative process. In the course of this discussion the theory of dynamic programming will be covered and some applications examined. To make dynamic programming more practical, one has to also develop a technique in applying it. To see what these techniques might be, we'll look at an example quite closely.

Definitions

So that we may proceed more clearly, some definitions must be set down.

Decision Process. If at a stage in a process one has a choice of activity he might apply to the system, there is a decision process. There are two types of decision processes, a single-stage process and a multi-stage process. In a single-stage process, an activity is applied to the system at only one point in the process. While in the multi-stage decision process are not necessarily the same at all points in the process.

Policy. Any choice of activity or sequence of activities is a policy. If the choice is such that the policy produces a maximum return, the policy is an optimal policy.

A Discussion of Decision Processes

Description of a Decision Process. Suppose there is available a certain quantity of a resource. This resource may be money, men, machines or materials such as water for agriculture or industry or fuel to operate machines. Now, if there are several ways of allocating these resources, a conflict of interests arises. The question now is how should these allocations be made considering all of the interests.

The inventory problem is an example of considering several interests. One must try to keep down the holding cost while at the same time keep down the shortage cost. Also, if the delivery cost is high one might want to ask for larger quantities delivered and fewer deliveries.

To determine what is the best allocation, one has to be able to measure the return resulting from each activity. In inventory this means that the total cost must be determined. A lower cost for one policy over another policy means that the first policy is preferred.

Construction of a Mathematical Model. It is possible to set down a decision process in mathematical form by determining a return function and the restraints. To write these functions, one must number each activity uniquely. That is, if there are N activities, they are numbered 1, 2, 3, ..., N. The way the activities are numbered is unimportant but once specified must be adhered to.

Return Functions show the relationship between the return from an activity and the amount of resource allocated. If x_i is the amount of resource allocated to activity i, then $g_i(x_i)$ is the return from the ith activity. If the activities are assumed to be independent of each other and the individual return functions are assumed to be additive, the total return function can be written as:

$$R(x_1, x_2, \ldots, x_n) = g_1(x_1) + g_2(x_2) + \ldots + g_n(x_n)$$

The ease in the maximization of this function depends on the constraints of the system.

<u>Constraints</u> are usually imposed on the system because of the resources available and because only positive resources can be assigned. The constraints take the form:

$$x_1 + x_2 + \ldots + x_n = x$$

and

$$x_i \geq 0, \quad i = 1, 2, \ldots, n$$

Other constraints of course must be added as necessary.

<u>Functional Equations.</u> These equations are the basis of dynamic programming. Suppose the return function is defined as:

$$R(x_1, x_2, \ldots, x_n) = g_1(x_1) + g_2(x_2) + \ldots + g_n(x_n)$$

with the restrictions:

$$x_1 + x_2 + \ldots + x_n = x$$

$$x_i \geq 0, \quad i = 1, 2, \ldots, n$$

To determine the optimal policy is to find the maximum return, and this optimizing function is defined as:

$$f_n(x_n) = \max_{\substack{x_1+x_2+\ldots+x_n=x \\ x_i \geq 0}} \left[R(x_1, x_2, \ldots, x_n) \right]$$

Let us observe:

$$f_n(x_n) = \max_{x_1+x_2+\ldots+x_n=x} \left[R(x_1, x_2, \ldots, x_n) \right]$$

$$= \max_{0 \leq x_n \leq x} \left[\max_{x_1+x_2+\ldots+x_{n-1}=x-x_n} \left[R(x_1, x_2, \ldots, x_n) \right] \right]$$

Then:

$$f_n(x_n) = \max_{0 \leq x_n \leq x} \left[\max_{x_1+x_2+\ldots x_{n-1}=x-x_n} \left[g_1(x_1) + g_2(x_2) + \ldots + g_n(x_n) \right] \right]$$

$$= \max_{0 \leq x_n \leq x} \left[g_n(x_n) + \max_{x_1+x_2+\ldots+x_{n-1}=x-x_n} \left[g_1(x_1) + \ldots + g_{n-1}(x_{n-1}) \right] \right]$$

Finally:

$$f_n(x) = \max_{0 \leq x_n \leq x} \left[g_n(x_n) + f_{n-1}(x - x_n) \right]$$

To use this function we must define:

$$f_1(x) = g_1(x_1)$$

Now $f_n(x)$ can be computed by an iterative process. This process will be demonstrated in an example later.

Principle of Optimality. The following is the definition given by Bellman and Dreyfus 3 . "An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."

The principle has already been established above. That is, $g_n(x_n)$ can be considered as the initial decision and $f_{n-1}(x - x_n)$ as the optimal policy constituting the remaining decisions.

An Application

To illustrate the use of dynamic programming, let us examine an inventory control problem as discussed by Bhatia and Garg [4]. In this problem the demand during each of the next n periods is known and it is desired that an optimal delivery schedule be determined. That is, to determine the least cost schedule. The following are defined:

r_i = Requirement for period i.

x_i = Quantity to be delivered at the beginning of period i.

S = Fixed cost per delivery.

λ = Inventory holding cost per unit per period.

$I(T,t)$ = Inventory holding cost from beginning of period T to end of period t.

$R(T,t)$ = Total cost to end of period t.

$f(t)$ = Minimized cost to end of period t.

The functional equation is written as:

$$R(T,t) = S + I(T,t) + f(T-1)$$

and :

$$f(t) = \min_{1 \leq T \leq t} \left[S + I(T,t) + f(T-1) \right]$$

with the following constraints:

$$\sum_{i=1}^{j} x_i \geq \sum_{i=1}^{j} r_i, \quad j = 1, 2, \ldots, n-1$$

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} r_i$$

$$f(0) = 0$$

The expression for $I(T,t)$ can be written in terms of x_T and r_i's as follows:

$$I(T,t) = \lambda \Big[(x_T - r_T) + (r_T)/2 + (x_T - r_T + r_{T+1}) + (r_{T+1})/2 + \ldots + (r_t)/2 \Big]$$

$$= \lambda \left[(t - T)x_T - \sum_{i=T}^{t} (t - i - \tfrac{1}{2}) r_i \right]$$

The functional equations are now in a form ready for computations. Let us now continue with the numerical example. The data and requirements are as follows [4].

Period	No. of Units
1	100
2	200
3	300
4	250
5	150

λ = \$.02 per unit per period

S = \$15.00 per delivery

The solution is given in Table III. It is noted that the costs are slightly different than those given by Bhatia and Garg [4]. Bhatia and Garg only considered the holding cost of inventory that is being held for depletion during a later period. However, while the total costs are slightly different, the optimal delivery schedule is the same.

TABLE III - Solution

t	T	X_1	X_2	X_3	X_4	X_5	f(T-1)	I(T,t)	S	R(t)	f(t)
		Delivery Schedule					**Costs**				
1	1	100					0	1	15	16*	$16.00
2	1	300	-				0	7	15	22*	$22.00
	2	100	200				16	2	15	33	
3	1	600	-	-			0	22	15	37*	$37.00
	2	100	500	-			16	11	15	42	
	3	300 #	- #	300			22	3	15	40	
4	1	850	-	-	-		0	39.50	15	54.50	
	2	100	750	-	-		16	23.50	15	54.50	
	3	300 #	- #	550	-		22	10.50	15	47.50*	$47.50
	4	600 #	- #	- #	250		37	2.50	15	54.50	
5	1	1000	-	-	-	-	0	53	15	68	
	2	100	900	-	-	-	16	34	15	65	
	3	300 #	- #	700	-	-	22	18	15	55*	$55.00
	4	600 #	- #	- #	400	-	37	7	15	59	
	5	300 #	- #	550 #	- #	150	47.50	1.50	15	64	

* Least Cost for all Periods Considered Through Period t.

\# Least Cost for all Periods Through Period T-1.

Comments

Dynamic programming has two advantages which are important in the solution to multi-delivery inventory systems. One is that it provides a procedure to solve allocation problems that could not be solved before. The second is that when optimizing, the number of necessary computations is reduced. When working with an n-stage inventory control problem such as the one that was illustrated, 2^{n-1} computations are necessary if all possible combinations are considered and $n(n+1)/2$ computations if a dynamic programming approach is used. However, the reduction in the number of computations is not realized until a problem of more than 5 stages is encountered [4].

The theory of dynamic programming is brief. It consists of a development of a functional equation and the principle of optimality. Later in another chapter it will be shown how dynamic programming can be used on the problem in this thesis in much the same way that Bhatia and Garg used it. Bhatia and Garg assumed a known demand while in solving the thesis problem an assumed distribution of demand will be used.

5. COMBINING DYNAMIC PROGRAMMING WITH MARKOV PROCESSES

A powerful mathematical procedure has been developed by Ronald Howard [11]. He combined dynamic programming with markov processes and developed a policy-iteration method which can be used to solve complex systems such as the inventory system that is being discussed. The concept of markov processes will be discussed first, then the concept of allowing alternatives will be added. This will lead into the necessity of using dynamic programming to reduce the number of computations.

Markov Processes

The basic concept of a markov process includes the idea of a state and the idea of a transition from one state to another state. If a system can take on one or more states that can be defined and if the system can change from one state to another, we have a markov process. To define a markov process specifically, one needs to know the probabilities in going from one state to another. A transition matrix is used in specifying all of the transition probabilities involved. Each element (P_{ij}) of this matrix is defined as the probability of going to state j if the system is in state i. The sum of all elements in a row must sum to 1 and the numerical value of each element is between zero and 1.

Let us now discuss an example. Ronald Howard [11] discusses the toymaker problem, so we will take a look at his discussion. In this problem there are two states, state 1 if the toymaker has a successful toy (selling his toy) and state 2 if he has an unsuccessful toy (not selling his toy). If he is in state 1, the probability of staying in state 1 is .50 and consequently, the probability of making a transition to state 2 is .50. If he is in state 2, the probability of making a transition to state 1 is .40 and, consequently, the probability of staying in state 2 is .60. That is $P_{11} = .50$, $P_{12} = .50$, $P_{21} = .40$, $P_{22} = .60$. In matrix form this is:

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} = \begin{bmatrix} .50 & .50 \\ .40 & .60 \end{bmatrix}$$

P is called the transition matrix.

Reward Matrix

One can associate a reward with each transition. That is, if a transition is made from state i to state j, a reward r_{ij} is associated with that transition. Now we can write these rewards in matrix form similar to the transition matrix. This matrix is called the reward matrix R. Suppose we again illustrate with the toymaker example used by Howard [11] where $r_{11} = 9$, $r_{12} = 3$, $r_{21} = 3$ and $r_{22} = -7$. Then the reward matrix is:

$$R = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} = \begin{bmatrix} 9 & 3 \\ 3 & -7 \end{bmatrix}$$

The negative values represent costs.

Suppose we have an N state system and are interested in determining the total reward for the next n periods given that the system is in state i. Using Howard's notation let $v_i(n)$ be defined as the total expected reward in the next n periods when starting in state i. The reward equation is:

$$v_i(n) = \sum_{j=1}^{N} p_{ij}\left[r_{ij} + v_j(n-1)\right]$$

where:

$i = 1, 2, \ldots, N$

$n = 1, 2, \ldots,$

The above equations can also be written as:

$$v_i(n) = \sum_{j=1}^{N} p_{ij}r_{ij} + \sum_{j=1}^{N} p_{ij}v_j(n-1)$$

if, $$q_i = \sum_{j=1}^{N} p_{ij}r_{ij}$$

then:

$$v_i(n) = q_i + \sum_{j=1}^{N} p_{ij}v_j(n-1)$$

The first term is the immediate reward while the second term is the reward for the remaining n-1 periods. As we can see $v_i(n)$ can be

Table IV. Total Expected Rewards

n	$v_1(n)$	$v_2(n)$
0	0	0
1	6	-3
2	7.5	-2.4
3	8.55	-1.44
4	9.555	-0.444
5	10.5555	0.5556

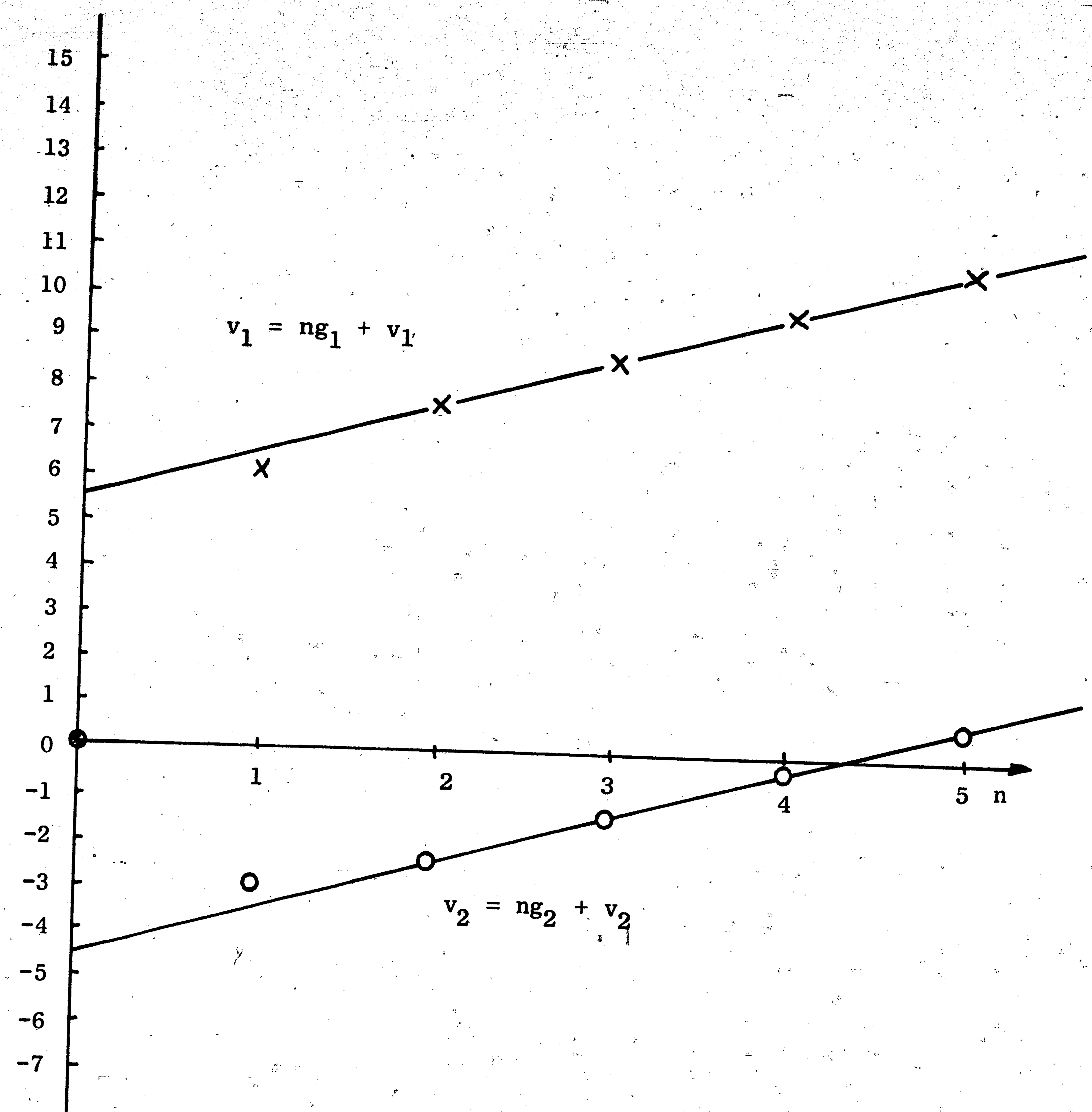

Figure 3

Total Expected Rewards

computed using an iterative process given a starting state $v_i(0)$. Again going to the toymaker example the numerical rewards can be computed as in Table IV.[11]. $v_1(0)$ and $v_2(0)$ are assumed to be zero.

One can see more clearly what happens with the total expected reward if the rewards in Table IV are plotted as shown in Figure 3 [11]. As the number of periods n increases, the expected reward approaches asymptotically a linear line. The equation of this line is:

$$v_i(n) = ng_i + v_i$$

where

g_i = slope of line

v_i = coordinate intercept of line

If the system is ergodic, then all g_i's are equal and we can write

$$v_i(n) = ng + v_i$$

We have now developed a method for computing rewards of a specific process. Now we will continue and introduce the idea of allowing alternatives and develop a method for determining the optimal policy.

Alternates

Suppose we again take the toymaker example to illustrate how the idea of working with alternatives will work. If the toymaker is in state 1, he can alter the transition probabilities and associated rewards by taking another action. If he does no advertising the transition probabilities and associated rewards are $[P_{1j}] = [.50\ .50]$ and $[r_{1j}] = [9\ \ 3]$ respectively. However, if he does advertise, the transition probabilities and associated rewards become $[P_{1j}] = [.80\ .20]$ and $[r_{1j}] = [4\ \ 4]$. Also, if the toymaker is in state 2, two alternatives are open to him, namely, do no research or do research. If he does no research the transition probabilities and associated rewards are $[P_{2j}] = [.40\ .60]$ and $[r_{2j}] = [3\ -7]$, and if he does research the transition probabilities and associated rewards become $[P_{2j}] = [.70\ .30]$ and $[r_{2j}] = [1\ -19]$ respectively.

Let us now use the superscript K to indicate the alternate chosen. Thus, we can write the optimal total reward equation as:

$$v_i(n) = \underset{K}{\text{Max}} \left[q_i^K + \sum_{j=1}^{N} P_{ij}^K v_j(n-1) \right]$$

To compute numerical values, $v_i(0)$ and $v_2(0)$ are again equal to zero. The maximum rewards for each period are given in Table V below. The notation $d_1(n)$ and $d_2(n)$ are defined as the alternative producing maximun reward for each state respectively.

Table V. Total Rewards

n	$v_1(n)$	$v_2(n)$	$d_1(n)$	$d_2(n)$
0	0	0		
1	6	-3	1	1
2	8.2	-1.7	2	2
3	10.22	0.23	2	2
4	12.222	2.223	2	2
.	.	.	.	.
.	.	.	.	.

We now have a way of evaluating alternative actions and determining the rewards of optimal decisions for n periods.

Next we will look at how we can evaluate and find optimal policies which means looking at a system where actions are evaluated for an infinite number of periods.

Policy-Iteration

The policy-iteration method as developed by Howard [11] consists of two parts, a value-determination operation and a policy-improvement routine. The value-determination operations computes the gain g and all the coordinate intercepts v_i of a given policy. And the policy-improvement routine uses the results of the value-determination operation to search for an improved policy. The procedure continues to cycle between these two parts until an optimal policy is determined.

Value-Determination Operation. As determined before the total expected reward can be written as:

$$v_i(n) = q_i + \sum_{j=1}^{N} P_{ij} v_j(n-1)$$

Also, we have found that in an ergodic process, $v_i(n)$ approaches a linear line of the form

$$v_i(n) = ng + v_i$$

If we substitute this linear function $v_i(n)$ into the total expected reward equation we have

$$ng + v_i = q_i + \sum_{j=1}^{N} P_{ij} \left[(n-1)\, g + v_j \right]$$

$$= q_i + (n-1)\, g \sum_{j=1}^{N} P_{ij} + \sum_{j=1}^{N} P_{ij} v_j$$

Since $\sum_{j=1}^{N} P_{ij} = 1$, this equation becomes:

$$g + v_i = q_i + \sum_{j=1}^{N} P_{ij} v_j \,, \qquad i = 1, 2, \ldots, N$$

We now have N simultaneous equations with N + 1 unknowns. To solve this set of equations we must set one v_i to zero. Suppose we set v_N to zero. Then we can solve for the values of all other v_i's and g and we are ready to go on to the policy-improvement routine.

Policy-Improvement Routine. As was seen before the test quantity is:

$$q_i^K + \sum_{j=1}^{N} P_{ij}^K v_j(n)$$

By substituting $v_j(n) = ng + v_j$ this test quantity becomes:

$$q_i^K + \sum_{j=1}^{N} P_{ij}^K \left[ng + v_j \right]$$

or

$$q_i^K + ng \sum_{j=1}^{N} P_{ij}^K + \sum_{j=1}^{N} P_{ij}^K v_j$$

Since the center term is constant it need not be carried along and the test quantity becomes:

$$q_i^K + \sum_{j=1}^{N} P_{ij}^K v_j$$

With the v_i's computed in the value-determination operation, we can now search for an improved policy by maximizing the test quantity.

$$\underset{K}{\text{Max}} \left[q_i^K + \sum_{j=1}^{N} P_{ij}^K v_j \right]$$

After a new improved policy has been found, we go back to the value-determination operation and compute new values for g and the v_i's. If the improved policy is the same as the one in the preceding cycle, an optimal policy has been found and the computation ceases.

An Example

Again we can use the toymaker problem to illustrate our procedure. In the value-determination operation the following simultaneous equations must be solved.

$$g + v_1 = q_1 + P_{11}v_1 + P_{12}v_2$$

$$g + v_2 = q_2 + P_{21}v_1 + P_{22}v_2$$

or

$$g + v_1 = 6 + .50v_1 + .50v_2$$

$$g + v_2 = -3 + .40v_1 + .60v_2$$

Setting $v_2 = 0$ and solving for v_1 and g yields $g = 1$ and $v_1 = 10$. Now we go into the policy-improvement routine. This procedure is illustrated in Table VI [11].

Table VI. Toymaker Policy-Improvement Routine

State	Alternative	Test Quantity $q_i^K + \sum_{j=1}^{N} P_{ij}^K v_j$
1	1	6 + 0.50(10) + 0.50(0) = 11
	2	4 + 0.80(10) + 0.20(0) = 12 ←
2	1	-3 + 0.40(10) + 0.60(0) = 1
	2	-5 + 0.70(10) + 0.30(0) = 2 ←

The arrows point to the alternative that gives an improved policy. We can now go back to the value-determination operation and compute new values for g, v_1 and v_2. The simultaneous equations are:

$$g + v_1 = 4 + .80v_1 + .20v_2$$

$$g + v_2 = -5 + .70v_1 + .30v_2$$

Again setting $v_2 = 0$ and solving for v_1 and g yields $g = 2$ and $v_1 = 10$. Note that the gain g has improved. Now we are ready to go back to the policy-improvement routine. Since v_1 and v_2 have the same value as before, the policy-improvement calculations will be the same as before in Table VI. Therefore, we have now completed the computations. The optimal policy is that if the toymaker is in state 1 (has a successful toy), he should choose alternative 2 (do advertising). If he is in state 2 (has no successful toy), he should choose alternative 2 (do research).

6. SOLUTION OF THE PROBLEM

We are now ready to work with the proposed inventory model. However, to make the solution more useful, let us expand the model from what was proposed in chapter 3. Let us increase the number of states to 15.

The Expanded Model

As before we will look ahead 5 weeks. Since the results of a decision made at the beginning of the month affects the system for 5 weeks, we must charge to this decision the rewards associated with this system through week 5. Also we do not charge to this decision until week 1 is completed. Week 1 for the present decision period is week 5 for the preceding decision period. The expanded model is shown in Figure 4.

It is further proposed that we allow the demand distribution function to take on 8 levels. As shown in Figure 4 the system can take on 15 states at the first and last stage of the time period we are looking at. The system can take on 22, 23, 26 and 21 states at stages 2, 3, 4 and 5 respectively.

Computation Procedure

The computation procedure includes computing the transition probabilities and the total immediate rewards for each state (costs in our problem) during the policy-improvement routine and computing the total expected rewards and gain of the system during the value-determination operation. This procedure is similar to the one used by Ken Larson [12] in his study of a single delivery inventory system.

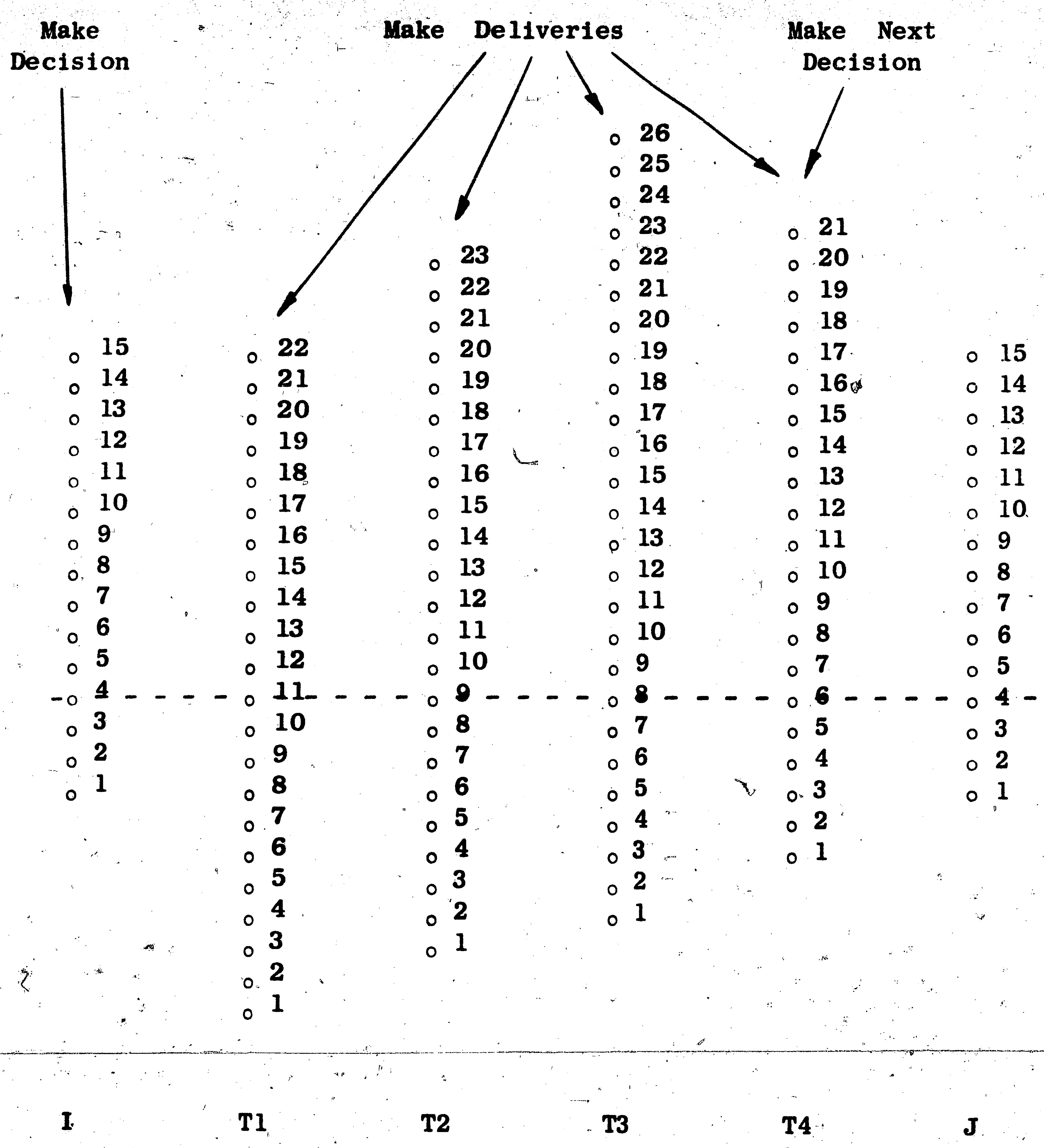

Figure 4

Expanded Inventory Model

We enter the iteration cycle in the value-determination operation. We assume a policy and then compute the transition matrix and immediate rewards for the given policy. We then go ahead and solve the 15 simultaneous equations below for g, v_1, v_2, ..., v_{14}; v_{15} is assumed to be zero.

$$g + v_i = q_i + \sum_{j=1}^{15} P_{ij} v_j \quad , i = 1, 2, \ldots, 15$$

After these equations are solved, we enter the policy-improvement routine.

Starting at each state of the system the transition probabilities are computed. The value of these probabilities will depend on the delivery policy chosen. So that not all possible policies need to be tried, dynamic programming will be used similar to the way Bhatia and Garg [4] used it. The difference being that they assumed a known demand while we will work with a discrete demand distribution. We can do this by working with 8 transitions at each stage and attaching the appropriate probability to each.

While we are computing transition probabilities, we can accumulate the immediate rewards associated with the improved policy. The rewards are negative in our system and include holding cost, ordering cost, delivery cost and shortage cost. The costs are computed as follows:

$$\text{Holding Cost} = \frac{(I_S + I_F)}{2} \times H \times P_S \times P_{SF} \,, \qquad \text{if } I_S > 0,\ I_F \geq 0$$

$$= \frac{I_S^2}{2(I_S - I_F)} \times H \times P_S \times P_{SF} \,, \qquad \text{if } I_S > 0,\ I_F < 0$$

$$= 0 \quad , \text{ if } I_S \leq 0$$

$$\text{Shortage Cost} = -I_F D \ , \text{ if } I_F < 0$$

$$= 0 \quad , \text{ if } I_F \geq 0$$

where:

S = Ordering Cost ($/order)

DC = Delivery Cost ($/delivery)

H = Holding Cost ($/unit/week)

D = Shortage Cost ($/unit/week)

I_S = Inventory State at Start of Week

I_F = Inventory State at End of Week

P_S = Probability of Being in Inventory State I_S

P_F = Probability of Making Transition to Inventory State I_F

As we proceed from week to week the costs are accumulated and the delivery alternative that costs least for each week chosen as the improved policy. The procedure is continued until all starting states have been considered. Then we proceed again to the value-determination operation. If the improved policy is identical to the preceding policy, we have completed computations and have found the optimal policy.

7. EVALUATION

Simulating Two Procedures

So that the above procedure can be evaluated, it can be simulated and the results compared with the results of a simulation using the procedure presently being used. In fact, we can go further and set up a factorial experiment which will also include variations of the costs being considered. The costs given in Table VII will be used. The ordering cost will not be included in the analysis because it has been found that an order is always placed, therefore, does not affect the economic analysis. In these simulations a normal weekly demand will be used. The mean is 4000 units/wk and the standard deviation is 1000 units.

The optimal policy, determined by using the proposed procedure and the likely costs given in Table VII, will be taken as the decision rules for the first simulation. The discrete demand distribution used in determining this optimal policy is given in Table VIII. We will note that this demand distribution is normal with the class marks put at the mean and at $\pm\,\sigma$, $\pm\,2\sigma$, and $\pm\,3\sigma$. The resulting decision rules are given in Table IX.

The decision rules for the second simulation are as below:

$$Q = FD + 1/3\ (DL - PL)$$

where:

$$FD = 4 \times \text{mean} = 16000 \text{ units/mo}$$

$$\text{Safety Stock} = 1.64 \times \sqrt{4}\,\sigma = 3280 \text{ units}$$

$$DL = \text{Safety Stock} + FD/4$$

$$= 3280 + 4000 = 7280 \text{ units}$$

TABLE VII. Costs Included In Analysis

	Low	Likely	High
Delivery Cost	$5.00	$12.00	$20.00
Holding Cost	.0058	.0192	.0337
Shortage Cost	.0385	.3846	1.3462

TABLE VIII. Discrete Demand Distribution

Demand	Probability of Demand
0	.0000
1000	.0062
2000	.0606
3000	.2417
4000	.3830
5000	.2417
6000	.0606
7000	.0062

TABLE IX. Decision Rules For First Simulation

If Inventory System Is In	Then Delivery Schedule Is			
	N_1	N_2	N_3	N_4
State 1 (-3,000)	9000	9000	4000	5000
State 2 (-2,000)	9000	8000	4000	5000
State 3 (-1,000)	9000	7000	4000	5000
State 4 (0)	9000	6000	4000	5000
State 5 (1,000)	9000	5000	4000	5000
State 6 (2,000)	8000	5000	4000	5000
State 7 (3,000)	7000	5000	4000	5000
State 8 (4,000)	6000	5000	4000	5000
State 9 (5,000)	5000	5000	4000	5000
State 10(6,000)	4000	5000	4000	5000
State 11(7,000)	3000	5000	4000	5000
State 12(8,000)	2000	5000	4000	5000
State 13(9,000)	1000	5000	4000	5000
State 14(10,000)	0	5000	4000	5000
State 15(11,000)	0	4000	4000	5000

And

$$N_1 = .10 \times Q$$

$$N_2 = .25 \times Q$$

$$N_3 = .25 \times Q$$

$$N_4 = .40 \times Q$$

Each simulation is made for a period of two years. The total cost will be used to make all comparisons. Table X lists the results of two replications for each cell in the factorial design.

Analysis of Variance

An analysis of variance can be made on the data given in Table X. This table includes data from two replications of the simulation. The results of the analysis of variance is given in Table XI. The F-statistics that are computed indicate that at the 95% confidence level, there is a significant difference between the total costs resulting from the two sets of decision rules, a significant difference resulting from the three shortage costs, and a significant difference resulting from the interaction between the sets of decision rules and the shortage costs. After studying the data in Table X more closely, we can note that as the shortage cost increases, the difference between the total costs resulting from the two sets of decision rules increases.

Difference of Two Means Test

To make a comparison between the two sets of decision rules, a Scheffe test [21] can be made on part of the data used in the analysis of variance. The data used is given in Table XIII and are the results of simulations using likely costs. The computations of the Scheffe test are shown in Figure 6. After comparing the statistic $A\sqrt{\hat{V}(C_M)}$ with the contrast C_M, we can state with 75% confidence that the use of the proposed decision rules will result in a lower total cost.

Another comparison that can be made is the Student's t-test [20]. For this test data other than the data given in Table X is used. More simulations can be made using the likely costs. This data is given in Table XII. Three 2-year simulations are made using the proposed decision rules and eight 2-year simulations are made using the decision rules presently used.

The t-statistic can now be computed as shown in Figure 5. After comparing this statistic with a table value, we can state with 95% confidence that the use of the proposed decision rules will result in a lower total cost.

The difference in confidence level of these two tests can be explained by noting that more data is used in the t-test. The t-test uses 11 pieces of data while Scheffe's test is restricted to 4 pieces of data.

TABLE X. Results of Factorial Experiment

Holding Cost	Shortage Cost	Model	Delivery Cost		
			$20.00	$12.00	$5.00
.0337	1.3462	1	$ 25377.73 32232.67	$ 24625.73 31488.67	$ 23967.73 30837.67
		2	184375.90 63682.50	183543.90 62850.50	182815.90 62122.50
	.3846	1	21830.35 24307.87	21078.35 23563.87	20420.35 22912.87
		2	61340.15 29996.22	60508.15 29164.22	59780.15 28436.22
	.0385	1	20553.58 21455.56	19801.58 20711.56	19143.58 20060.56
		2	17056.92 17871.83	16224.92 17039.83	15496.92 16311.83
.0192	1.3462	1	17490.26 24023.91	16738.27 23279.91	16080.27 22628.91
		2	180137.32 57554.13	179305.32 56722.13	178577.32 55994.13
	.3846	1	13942.89 16099.11	13190.89 15355.11	12532.89 14704.11
		2	57101.64 23867.86	56269.64 23035.86	55541.64 22307.87
	.0385	1	12666.12 13246.81	11914.12 12502.81	11256.12 11851.81
		2	12818.41 11743.49	11986.41 10911.50	11258.41 10183.50
.0058	1.3462	1	10201.16 16437.90	9449.17 15693.91	8791.17 15042.91
		2	176220.38 51890.68	175388.38 51058.68	174660.38 50330.68
	.3846	1	6653.79 8513.09	5901.79 7769.09	5243.79 7118.09
		2	53184.66 18204.42	52352.66 17372.43	51624.66 16644.43
	.0385	1	5377.01 5660.79	4625.01 4916.79	3967.01 4265.79
		2	8901.45 6080.05	8069.45 5248.05	7341.45 4520.05

TABLE XI. Analysis Of Variance

SOURCE	DF	SS	MS
Total	107	231153559799.8996	
D	2	39575540.7459	19787770.3730
H	2	2914917281.2762	1457458640.6381
S	2	63186342691.6986	31593171345.8493
M	1	44947431817.7547	44947431817.7547
DH	4	2660.0203	665.0051
DS	4	2665.5803	666.3951
DM	2	113969.7036	56984.8518
HS	4	2731.6435	682.9109
HM	2	137538974.2847	68769487.1424
SM	2	47120645146.8205	23560322573.4103
DHS	8	5439.5770	679.9471
DHM	4	2640.4011	660.1003
DSM	4	2641.2555	660.3139
HSM	4	2708.8044	677.2011
DHSM	8	5461.5154	682.6894
RESIDUAL	54	72806967428.8179	1348277174.6077

$F_H = 1.081$, $F_{.95}(2,54) = 3.17$

$F_S = 23.432$, $F_{.95}(2,54) = 3.17$

$F_M = 33.337$, $F_{.95}(1,54) = 4.02$

$F_{SM} = 17.474$, $F_{.95}(2,54) = 3.17$

Table XII. Data For Student's t-test

X	Y
16105.72	65165.06
15147.87	17065.55
13465.58	16102.85
	61839.03
	24648.78
	19436.71
	54644.39
	29192.30

$\bar{X} = 14,906.39$ $\bar{Y} = 36011.83$

$$t = (\bar{y} - \bar{x}) \Big/ \sqrt{\frac{\hat{\sigma}_y^2}{n_y} + \frac{\hat{\sigma}_x^2}{n_x}} = 2.835$$

$$\nu = \frac{\left(\frac{\hat{\sigma}_y^2}{n_y} + \frac{\hat{\sigma}_x^2}{n_x}\right)^2}{\frac{\left(\frac{\hat{\sigma}_y^2}{n_y}\right)^2}{n_y + 1} + \frac{\left(\frac{\hat{\sigma}_x^2}{n_x}\right)^2}{n_x + 1}} - 2 = 9.2 - 2 = 7.2$$

Where:

$\bar{x}$ and $\bar{y}$ are sample means of data in Table XII.

$\hat{\sigma}_x$ and $\hat{\sigma}_y$ are unbiased estimates of the deviation of the data.

n_x and n_y are the number in each sample.

ν is the number of degrees of freedom.

From t-Tables:

$$t_{.95}(\nu) = t_{.95}(7) = 2.365$$

Figure 5. Computation of t-Statistic

Table XIII. Data For Scheffe Test

X	Y
13190.89	56269.64
15355.11	23035.86
$\bar{X} = 14273.00$	$\bar{Y} = 39652.75$

$$C_M = \sum_{i=1}^{t} C_{iM} T_i = 50{,}759.50$$

$$A = \sqrt{(t-1) F_{.75}(\nu_1, \nu_2)} = 0.679$$

$$\hat{V}(C_M) = \sum_{i=1}^{t} C_{iM}^2 \hat{V}(T_i) = 54.0 \times 10^8$$

$$\hat{V}(T_i) = n_i s^2$$

$$A\sqrt{\hat{V}(C_M)} = 49{,}800$$

Where:

C_M = Contrast between models.

C_{iM} = Coefficients in computing contrast.

T_i = Sum of n_i observations.

$\hat{V}(C_M)$ = Estimated variance of C_M.

t = Number of models compared.

n_i = Number of replicates in each cell.

s^2 = Estimate of experimental error.

Figure 6. Computation Of The Statistic $A\sqrt{\hat{V}(C_M)}$

8. CONCLUSION

We have discussed a method for solving a multi-delivery inventory system by using dynamic programming and markov processes. Dynamic programming is used as a search technique in finding optimal solutions while markov processes is used so that the probabilistic nature of the problem can be taken into consideration.

Dynamic programming has been very helpful for it indeed has reduced the computation time. Markov processes has given us a way to handle a probabilistic demand, however, it has also increased the computation time and negated some of the progress made with dynamic programming. It is in this area where further study might be made. It might be suggested that rather than working with the demand distribution during the transition stage in the problem, one should use the mean of the distribution for computing total costs and use the standard deviation for computing safety stock. This process would again cut down the computing time. This process would not be as pure, but the savings in time may make it attractive especially when the demand distribution is not known with high confidence anyway.

BIBLIOGRAPHY

1. Bahary, E. S., Research Engineer, Western Electric Company, Inc., private communication.

2. Bellman, R. E., Dynamic Programming, Princeton University Press, Princeton, New Jersey, 1957.

3. Bellman, R. E., and S. E. Dreyfus, Applied Dynamic Programming, Princeton University Press, Princeton, New Jersey, 1962.

4. Bhatia, A. and A. Garg, "Application of Dynamic Programming to a Class of Problems in Inventory Control," The Journal of Industrial Engineering, Vol. XI, No. 6, Nov.-Dec., 1960.

5. Bowman, E. H., and R. B. Fetter, Analysis for Production Management, Richard D. Irwin, Inc., Homewood, Illinois, 1961, Chapters 5, 9, and 10.

6. Buffa, E. S., Models for Production and Operations Management, John Wiley and Sons, Inc., New York, 1963, Chapters 15, 16, and 17.

7. Eilon, S., Elements of Production and Control, The Macmillan Co., New York, 1962, Chapter 17.

8. Elmaghraby, S. E., Research Leader, Western Electric Company, Inc., private communication.

9. Goode, H. H. and R. E. Machol, Systems Engineering, McGraw-Hill Book Company, Inc., New York, 1957, Chapter 4.

10. Hall, A. D., A Method for Systems Engineering, D. Van Nostrand Company, Inc., Princeton, New Jersey, 1962, Chapter 14.

11. Howard, R. A., Dynamic Programming and Markov Processes, The MIT Press, Cambridge, Massachusetts, 1960.

12. Larson, K. E., Research Engineer, Western Electric Company, Inc., private communication.

13. Littschwager, J. M., "Dynamic Programming in the Solution of a Multistage Reliability Problem," The Journal of Industrial Engineering, Vol. XV, No. 4, July-August, 1964.

14. Maffei, R. B., "Planning Advertising Expenditures by Dynamic Programming Method," Management Technology, Vol. 1, No. 2, Dec., 1960.

15. McGarrah, R. E., Production and Logistics Management, John Wiley and Sons, Inc., New York, 1963, Chapter. 7.

16. Saaty, T. L., Mathematical Methods of Operations Research, McGraw-Hill Book Company, Inc., New York, 1959, Chapters 5, 8, and 9.

17. Sasieni, M., A. Yaspan and L. Friedman, Operations Research-Methods and Problems, John Wiley and Sons, Inc., New York, 1959, Chapter 10.

18. Wagner, H. M., Statistical Management of Inventory Systems, John Wiley and Sons, New York, 1962.

19. Wagner, H. M., and T. M. Whitin, "Dynamic Version of the Economic Lot Size Model," Management Science, Vol. 5, No. 1, October, 1958.

20. Hoel, P. G., Introduction to Mathematical Statistics, John Wiley and Sons, New York, 1962, Chapter 11.

21. Ostle, B., Statistics In Research, The Iowa State University Press, Ames, Iowa, 1963, Chaptor 11.

VITA

PERSONAL HISTORY:

Name:	Fred Walter Ehrman
Birth Place:	Eureka, South Dakota
Birth Date:	December 9, 1929
Parents:	Fred and Margaret Ehrman
Wife:	Betty Lou Katherine Ehrman
Children:	Fred David and Kathleen Louise Ehrman

INSTITUTIONS ATTENDED:

Eureka High School Eureka, South Dakota	Graduated	1948
South Dakota School of Mines and Technology B. S. M. E.	Graduated	1958
Newark College of Engineering M. S. M. E.	Graduated	1961
Lehigh University Candidate for M. S. I. E.		1965

HONORS:

Sigma Tau Fraternity

Who's Who Among Students in American Colleges and Universities

PROFESSIONAL EXPERIENCE:

Engineer, Development Engineer, Research Engineer:
Western Electric Company, Incorporated

PROFESSIONAL ORGANIZATIONS:

Member, American Society of Mechanical Engineers.

Member, American Management Association